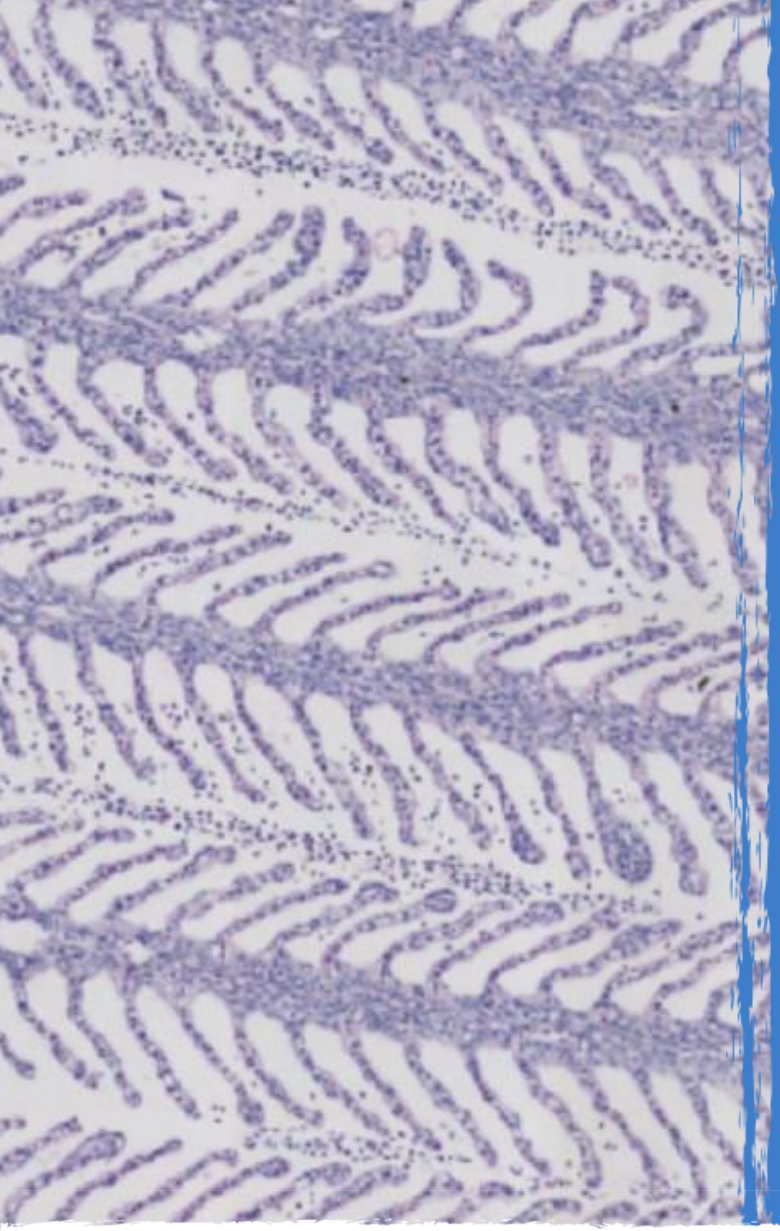


SVIP SRUC UNIVERSITY of STIRLING

Artificial Intelligence for Computer-Assisted Diagnosis of Hyperplasia in Atlantic Salmon Gill Histology Images

Alexander (Sandy) Carmichael

Johanna L. Baily, Aaron Reeves, Gabriela Ochoa, Annette S. Boerlage, Jimmy Turnbull, George Gunn, Rosa Allshire, and Deepayan Bhowmik

BE THE DIFFERENCE

The Problem – Analysing Epithelial Hyperplasia

Quantifying = % of epithelial hyperplasia

Advantages:

- Multiple samples: ranking and comparing
- Single sample: background information that can support diagnosis

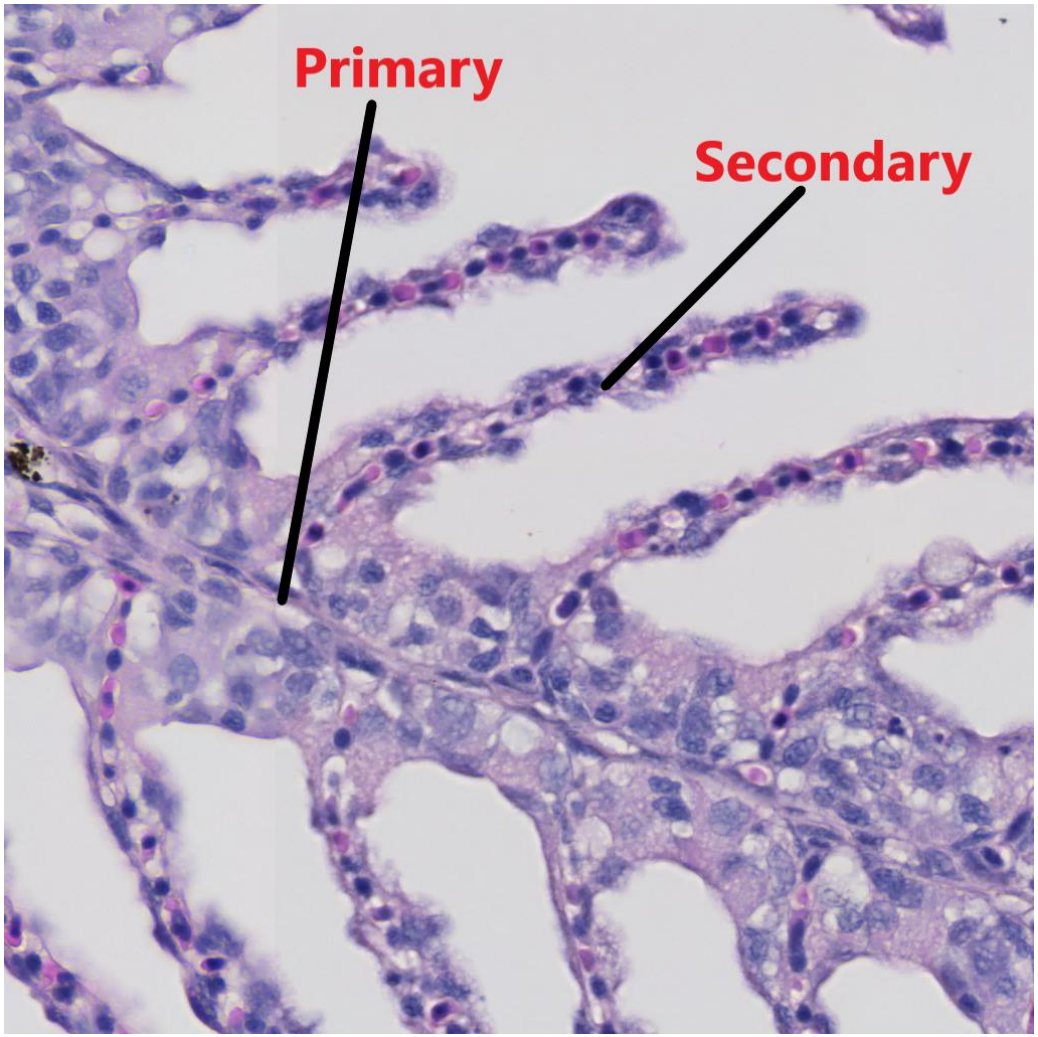
Disadvantages:

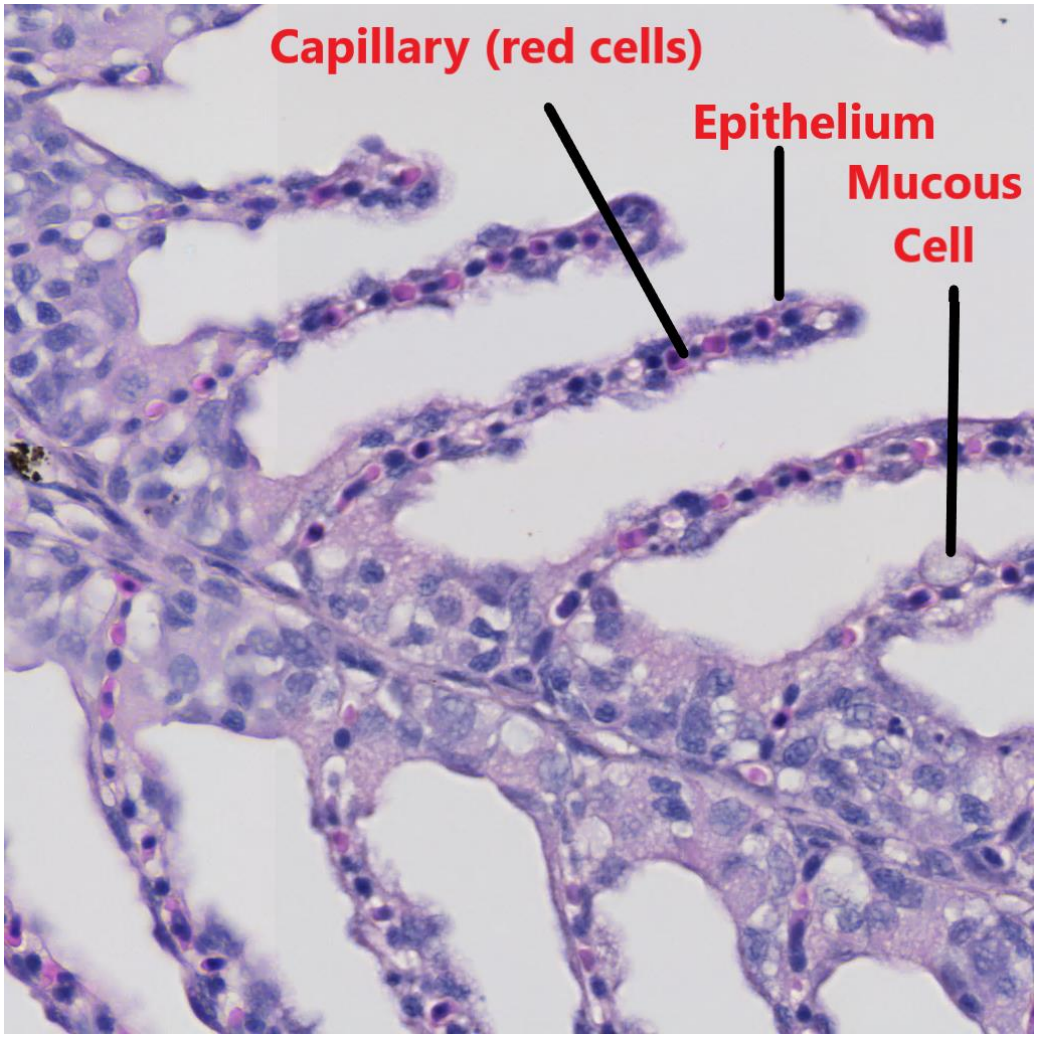
- Repetitive, time-consuming, costly process → worth the effort?
- Repeatability between individuals may be low.

Project Aims

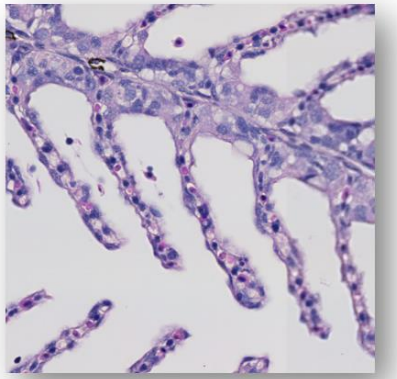
Develop a **computer-assisted diagnosis (CAD)** tool to automatically process Hematoxylin and Eosin (H&E) stained whole-slide images (WSIs) of gills.

- Support histopathologist with metrics
  - **severity** (%)
  - **distribution** (focal vs diffuse)
- Support histopathologist with target areas to look at
  - heatmap visualisations
- Other characteristics
  - 100% repeatability
  - Not a “black box” tool

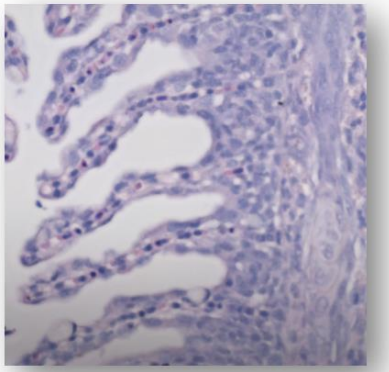

Primary
Secondary


Capillary (red cells)
Epithelium
Mucous
Cell

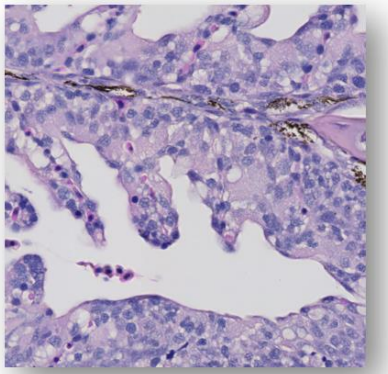
Epithelial Hyperplasia in Gills



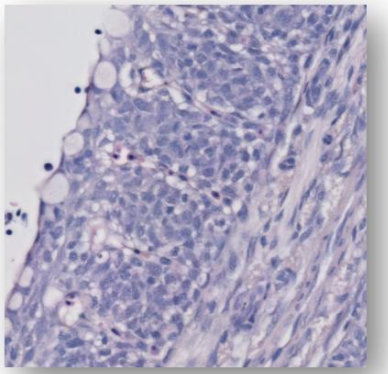
Normal



Mild



Moderate

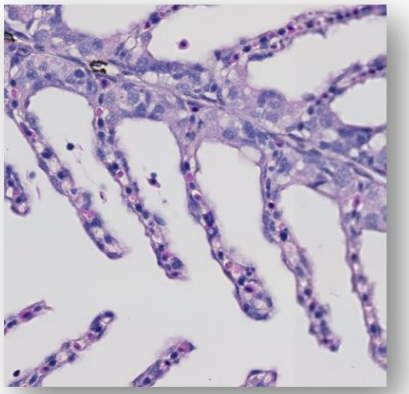


Severe

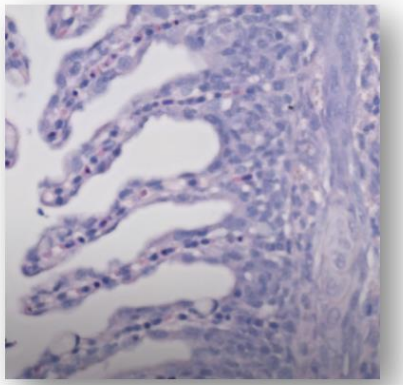
Histology Image Processing - Challenges

- **Large** image size.
- Lack of **labelled** data.
- Varying **colour** and **magnification**.
- **Stain intensity** varies from slide to slide.

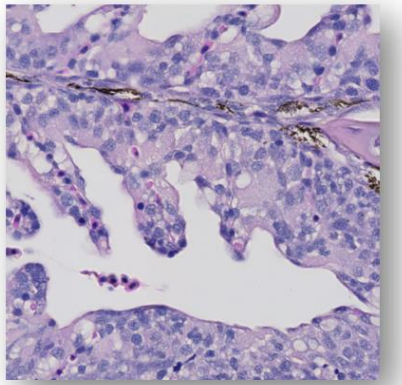
Epithelial Hyperplasia in Gills



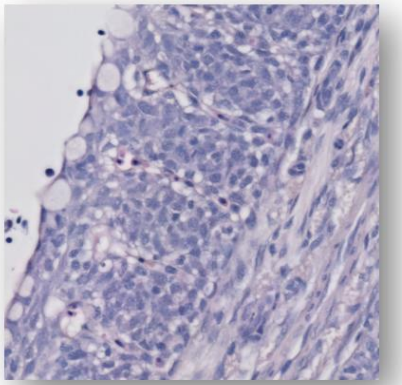
Normal



Mild



Moderate



Severe

Image Tiling

Image Tiling

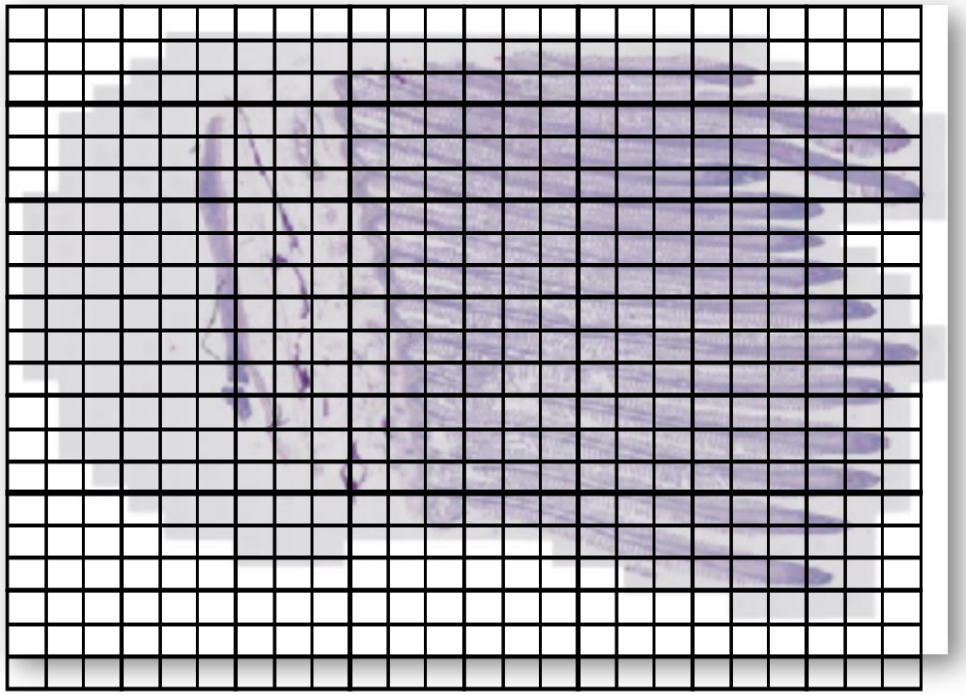
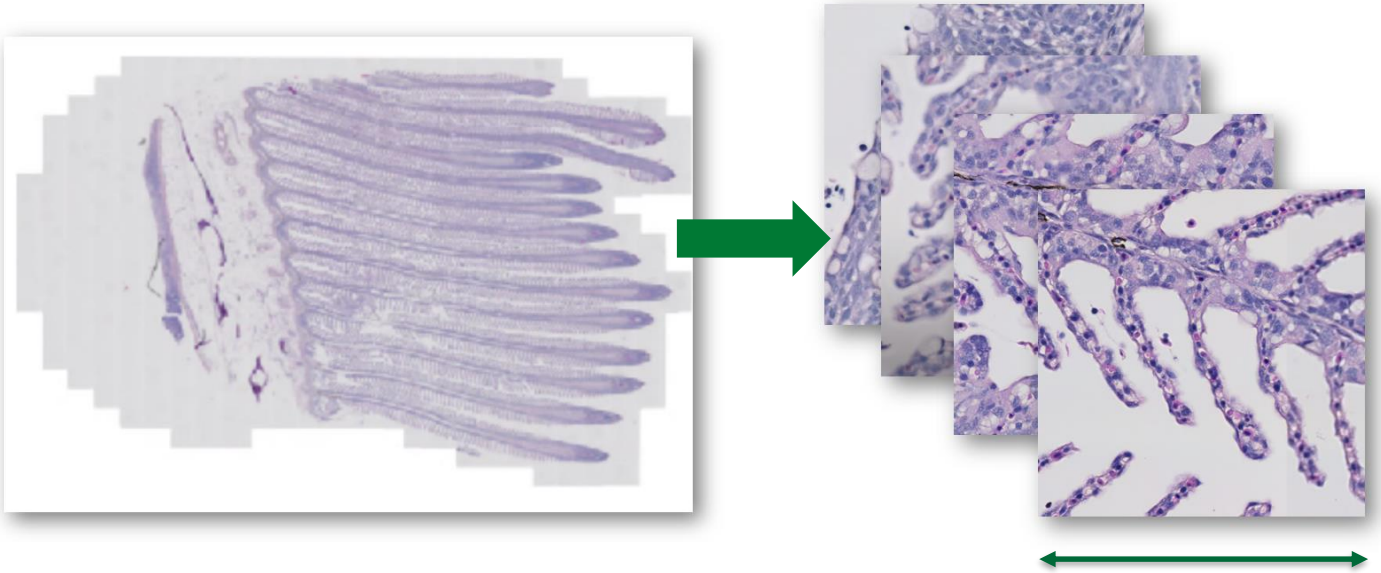


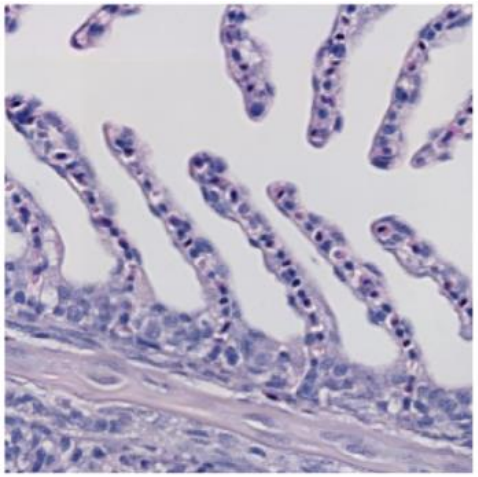
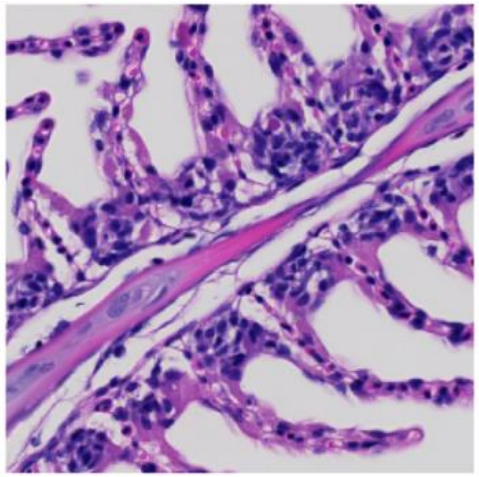
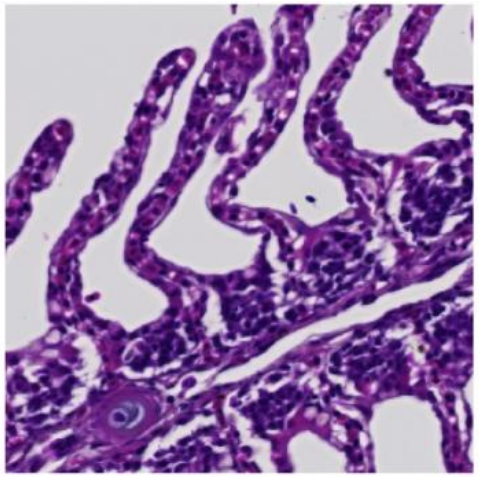
Image Tiling



Divide WSI into regions of interest (ROIs)

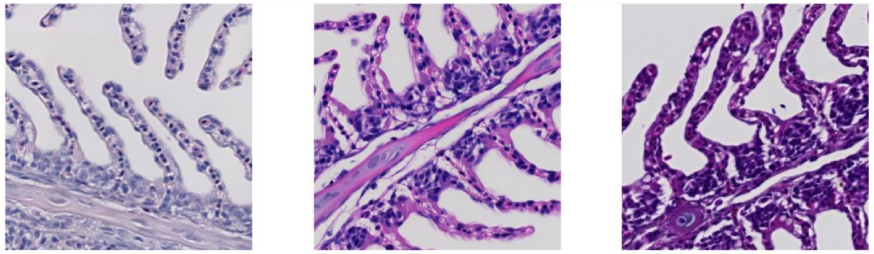
1024 pixels

Why stain normalisation?

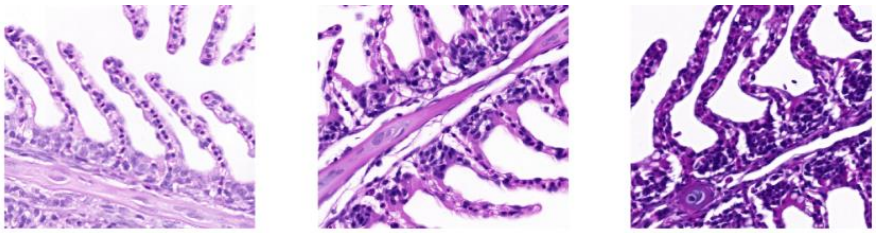




- **Stain intensity and colour** can vary between WSIs.

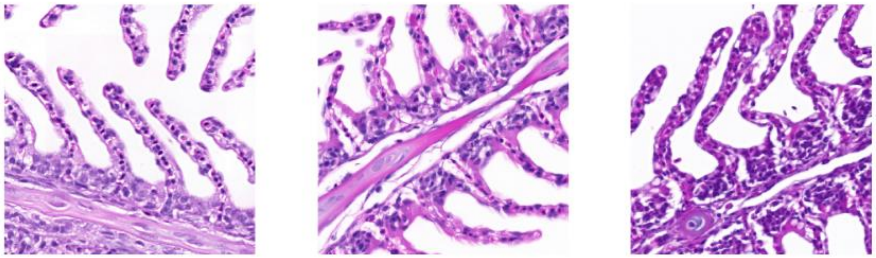
Tile



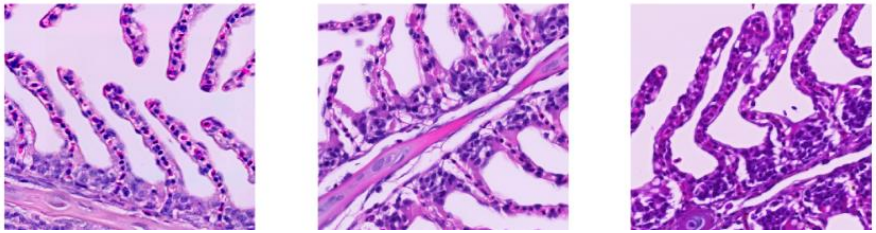
Vahadane



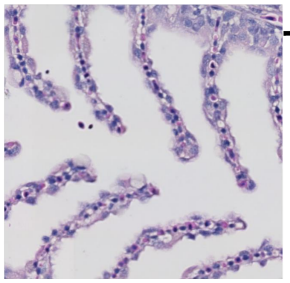
Macenko



Reinhard



WSI Analysis Framework



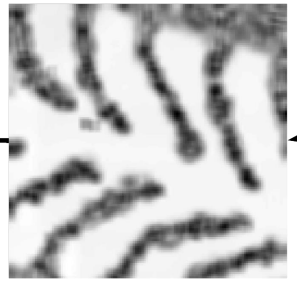
Stain Normalisation

Isolate Channel

Resize Image

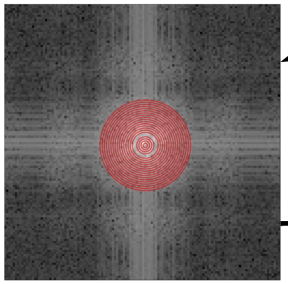
Apply 5x5 Blur

Normalise

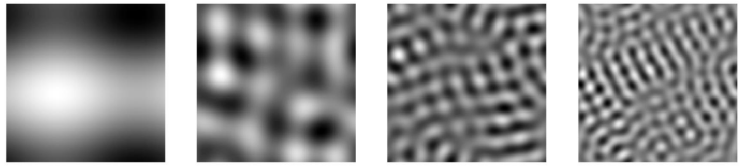


Convert to Frequency Domain

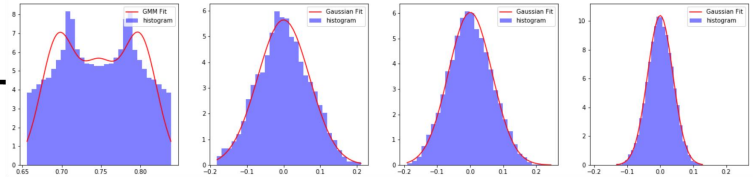
Get Boundary Scales



Get Subband Images

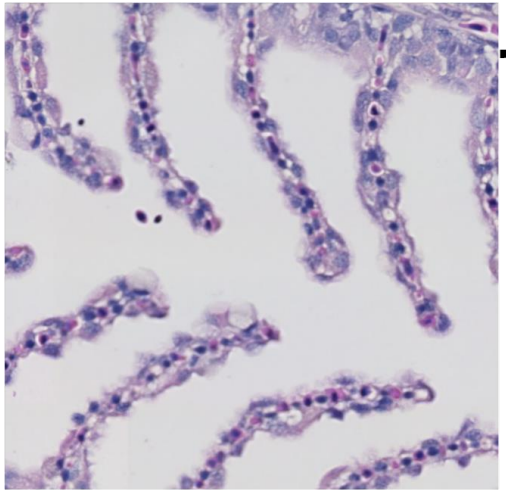


Calculate Subband Statistics



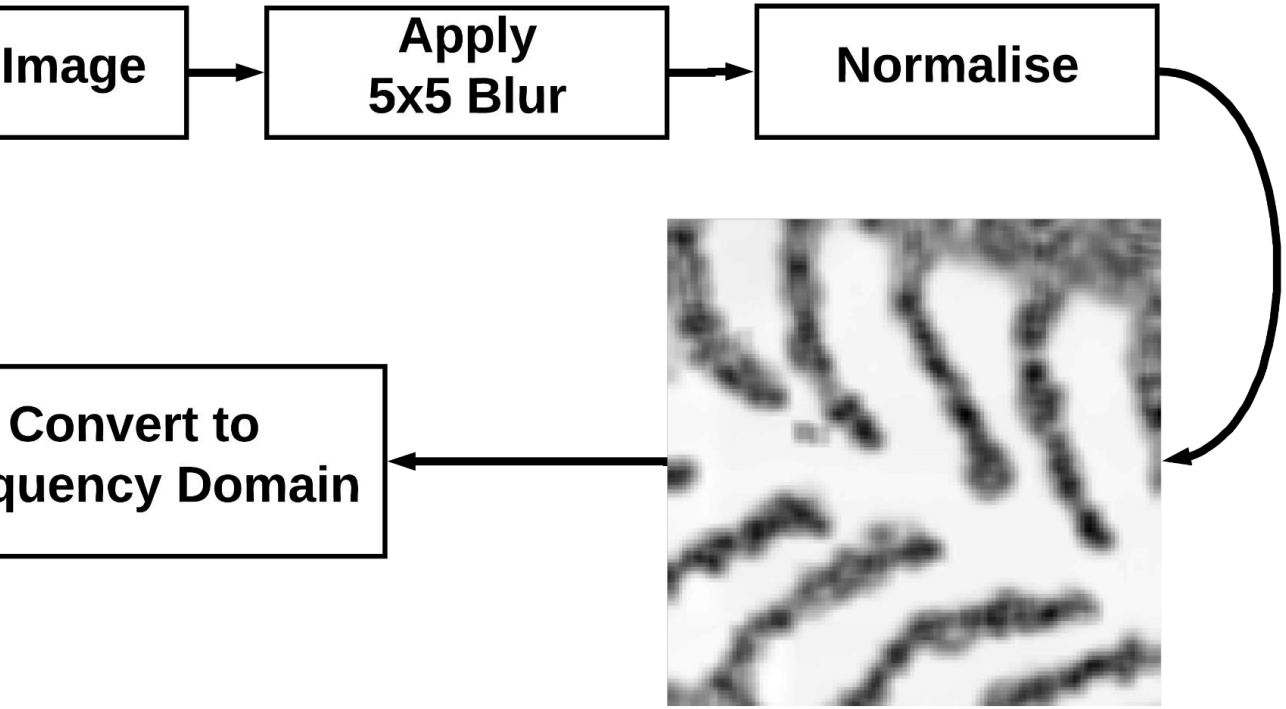
Concatenate Features into Feature Vector

Machine Learning Model (SVM, NN, etc.)

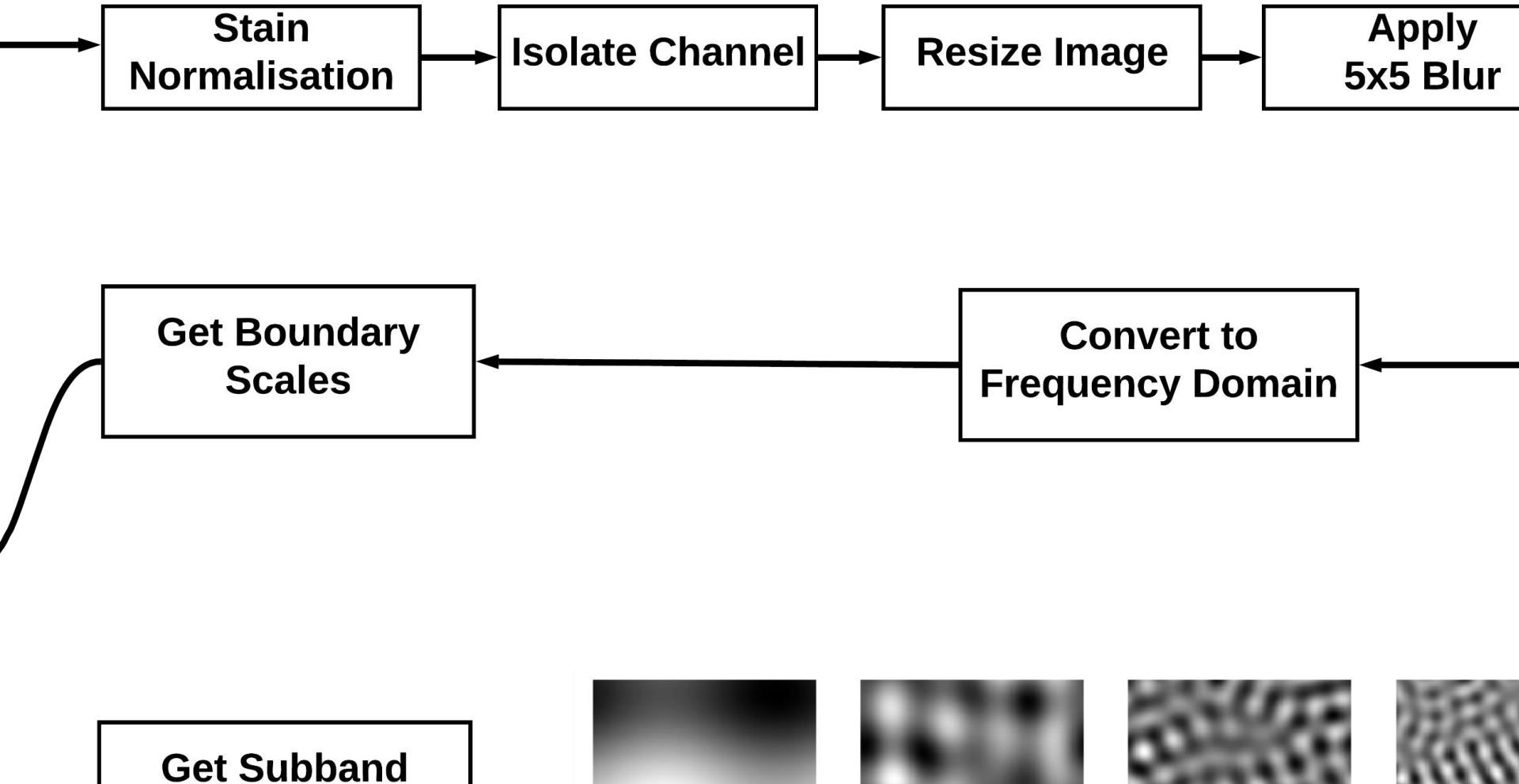


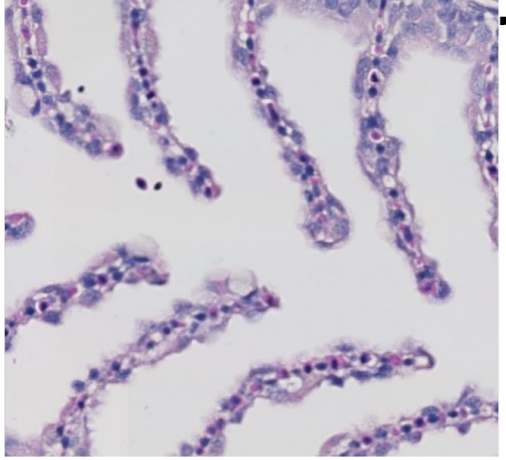
Stain Normalisation

Isolate



Calculate Subband


Stain Normalisation
Isolate Channel
Resize Image
Apply 5x5 Blur
Get Boundary Scales
Convert to Frequency Domain
Get Subband



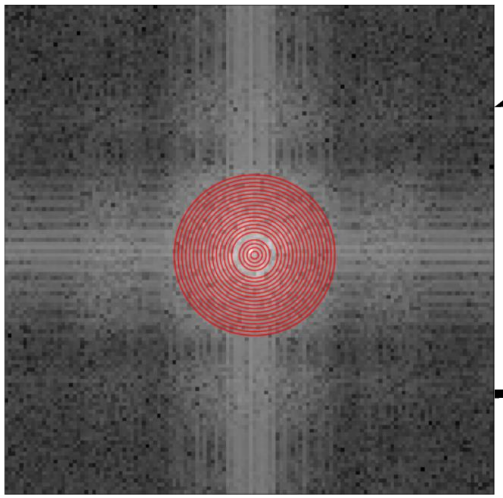
Stain Normalisation

Get Boundary Scales

$$\mathcal{W}_f^{\mathcal{E}\mathcal{L}\mathcal{P}}(n,\mathbf{x}) = \mathcal{F}_2^*\left(\mathcal{F}_2(f)(\omega)\overline{\mathcal{F}_2\left(\psi_n\right)(\omega)}\right)$$

$$\widetilde{\mathcal{F}}_P(|\omega|) = \frac{1}{N_\theta}\sum_{i=0}^{N_\theta-1}\left|\mathcal{F}_P(f)\left(\theta_i,|\omega|\right)\right|.$$

$$\mathcal{B} = \left\{\phi_1(\mathbf{x}), \{\psi_n(\mathbf{x})\}_{n=1}^{N-1}\right\}$$

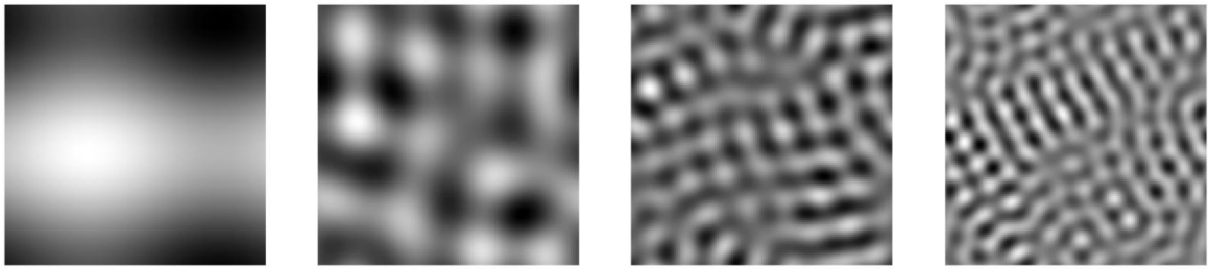


Get Subband Images

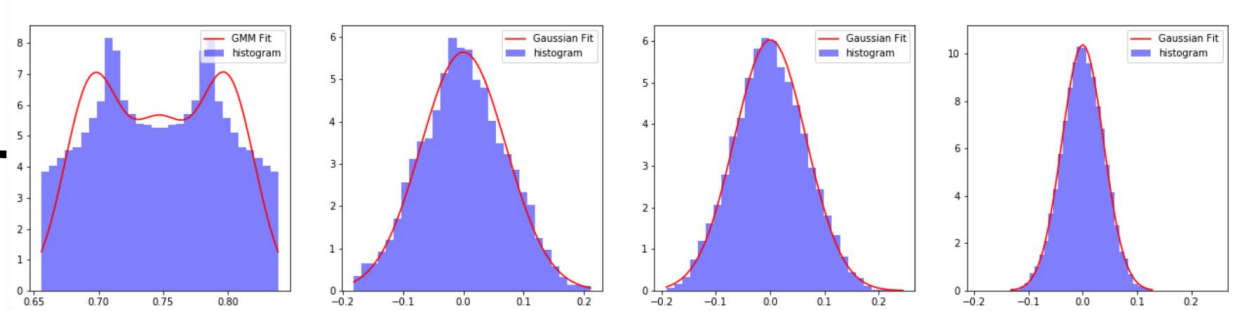
Machine Learning Model (SVM, NN, etc.)

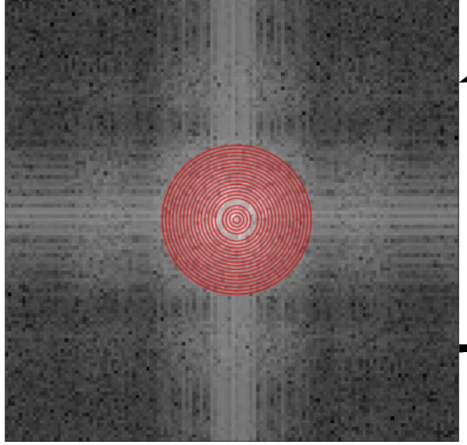
Concatenate Features into

Convert to Frequency Domain

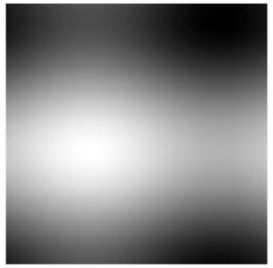


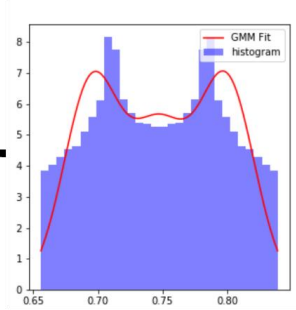
Calculate Subband Statistics

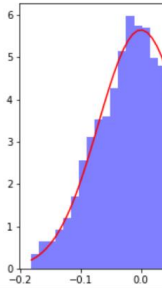




Get Subband Images

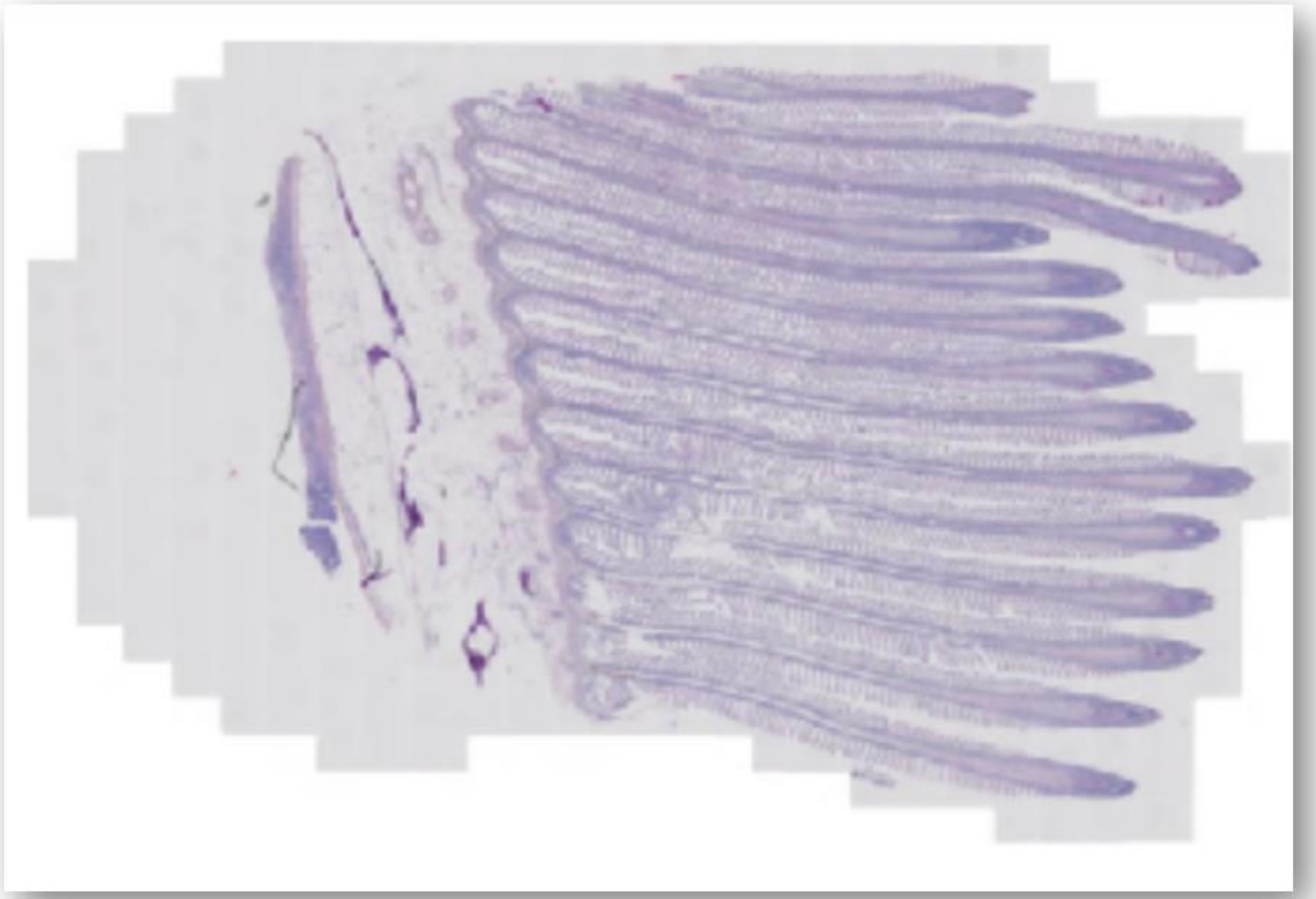


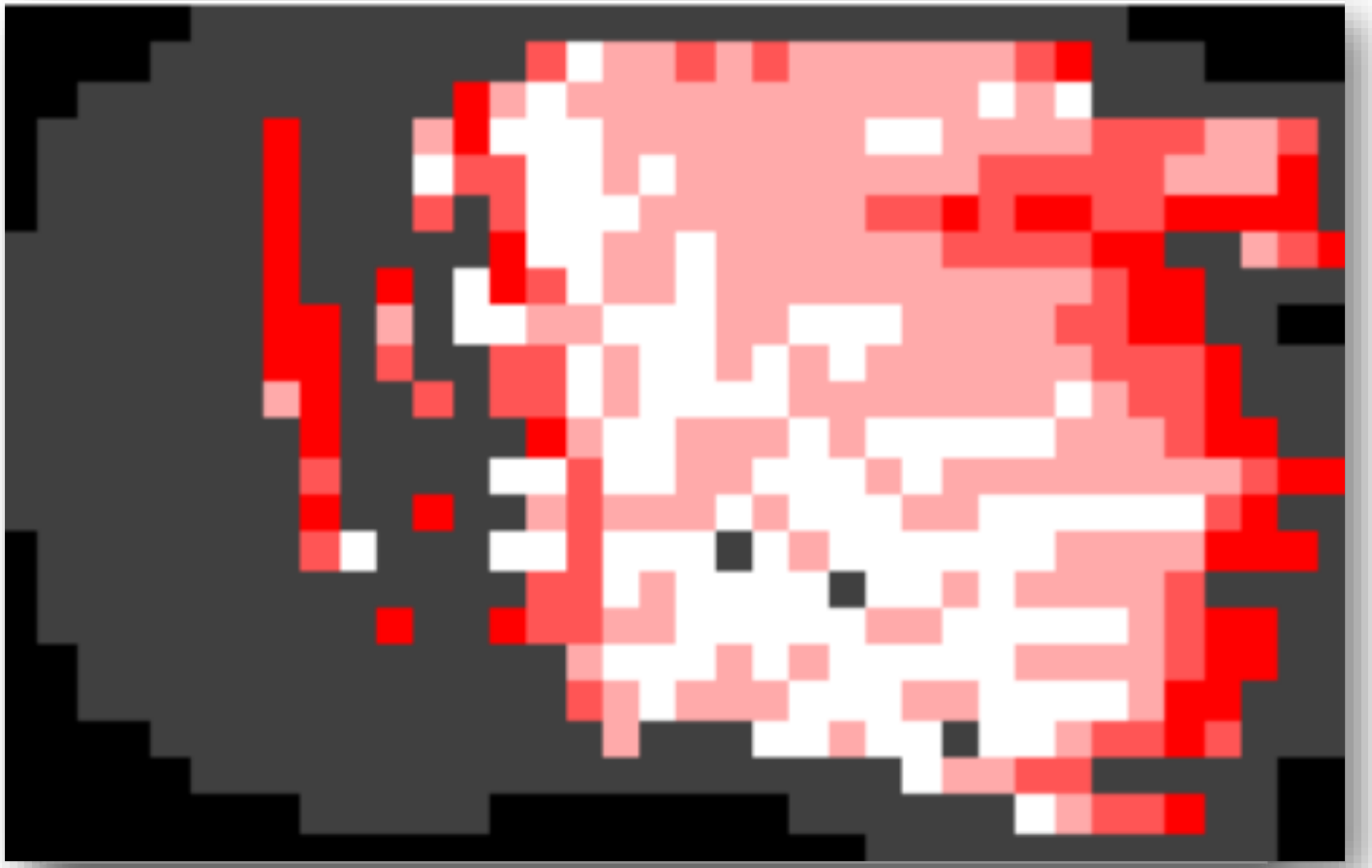


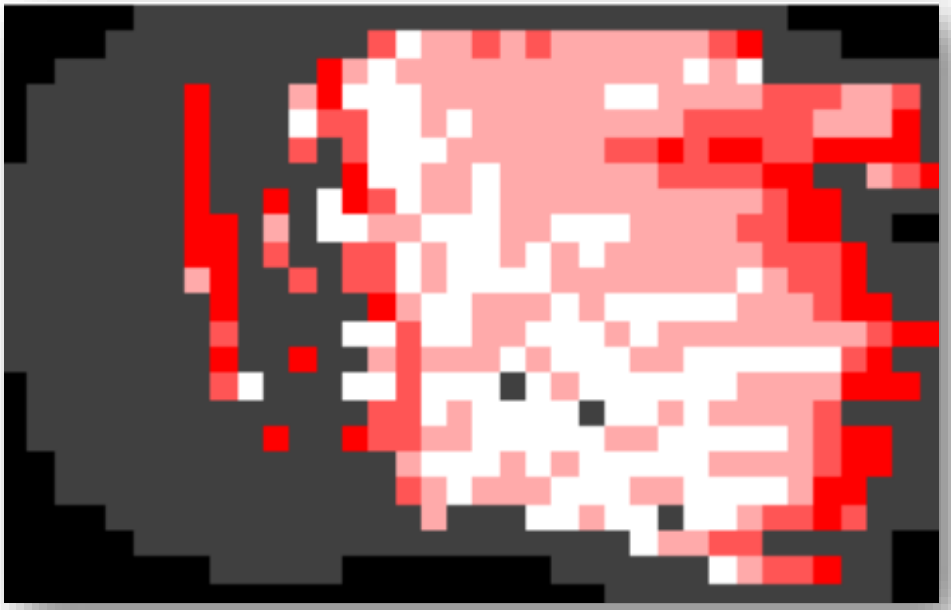


Concatenate Features into Feature Vector

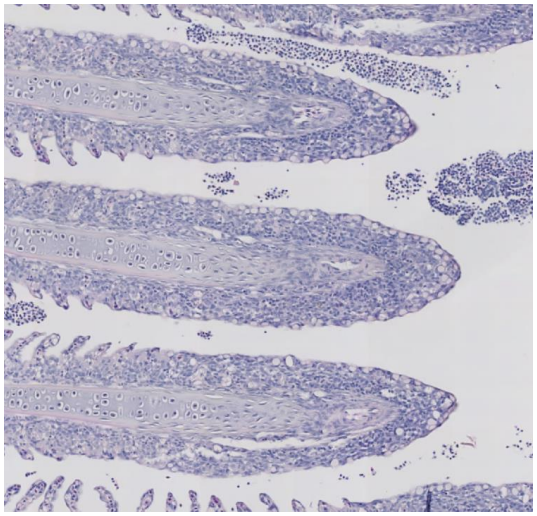
Machine Learning Model (SVM, NN, etc.)

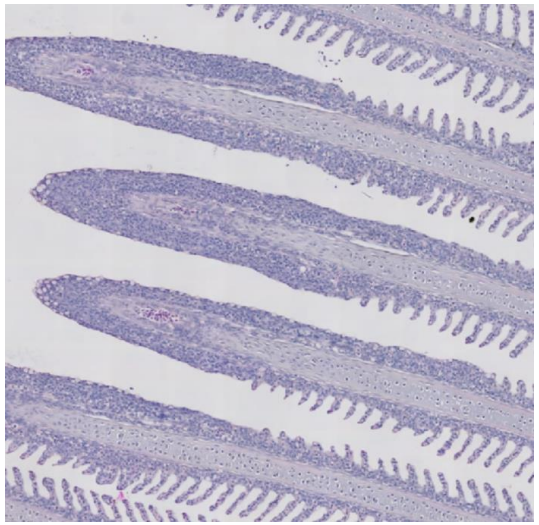


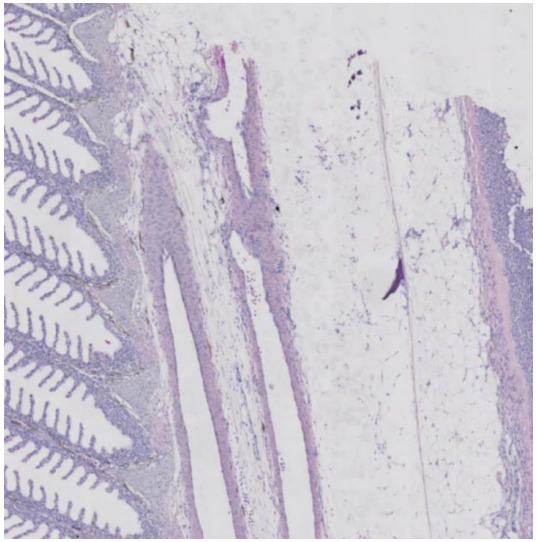


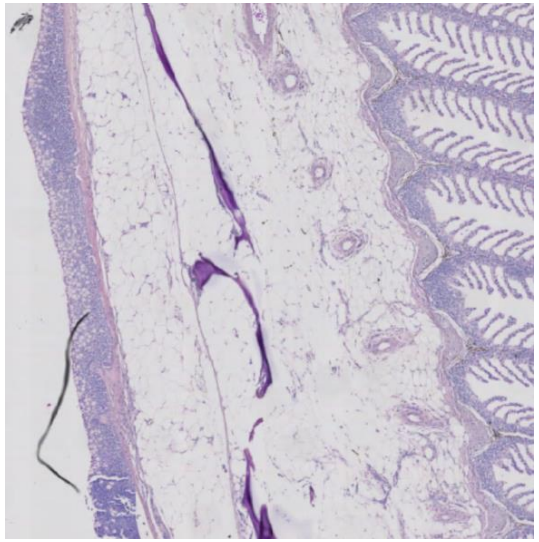


How can we make our approach aware of context?









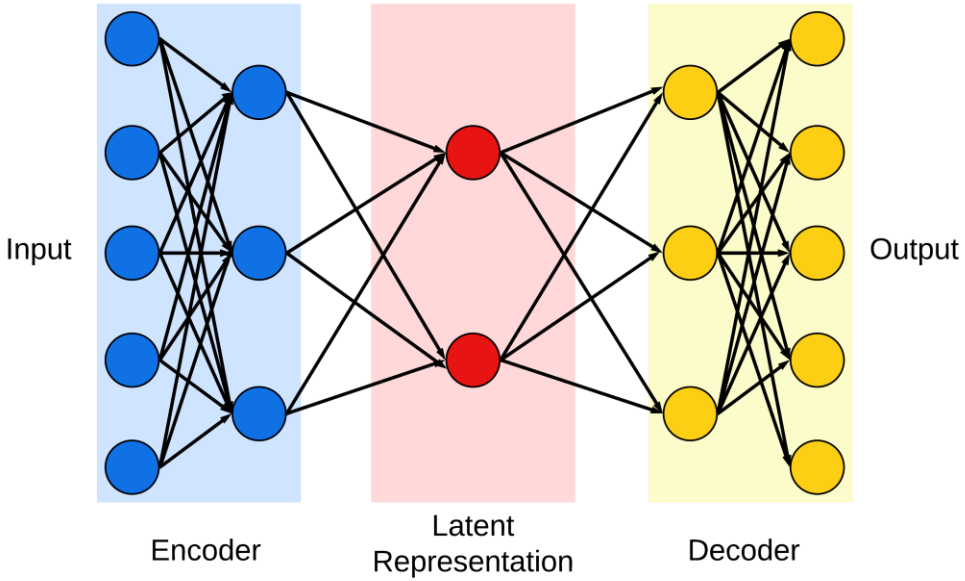
Some machine learning concepts…

- **Supervised learning** uses **labelled** data to train algorithms to make predictions, while **unsupervised learning** uses **unlabelled** data to uncover patterns or structures within the data.

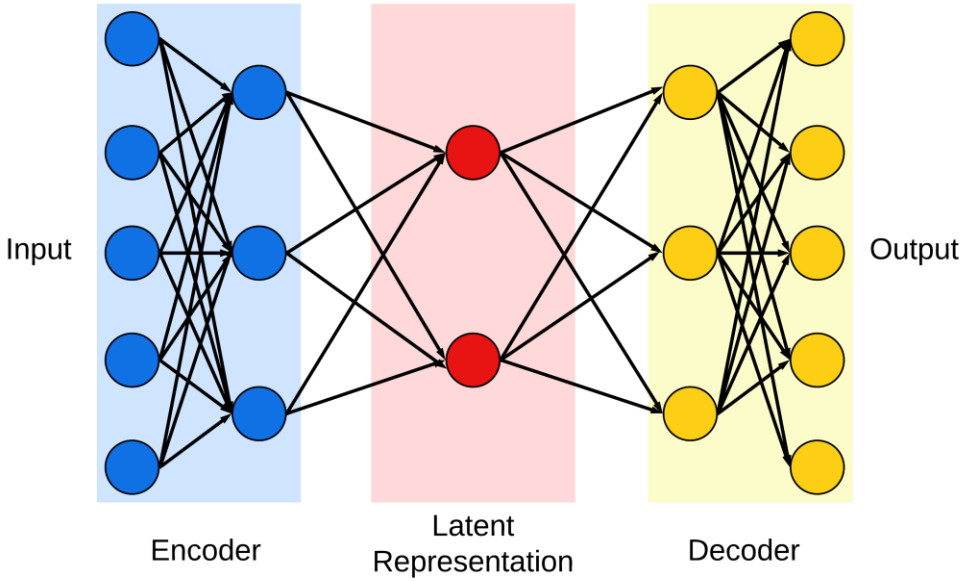
- So far, we have only used **supervised learning.** If we use unsupervised learning, we can train a model to identify salient **Regions of Interest (ROI)**.

- In other applications, this is known as **anomaly detection**.

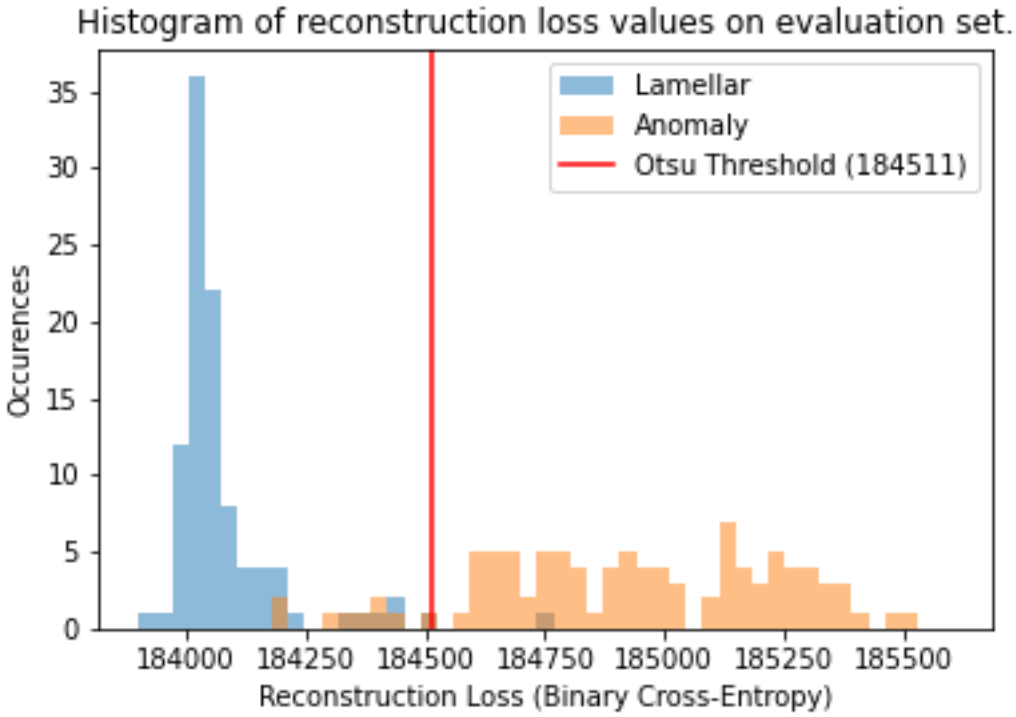
Reconstruction models try to compress, then recreate images. By only training it on lamellar tissue, we make the model learn about gill tissue structure.



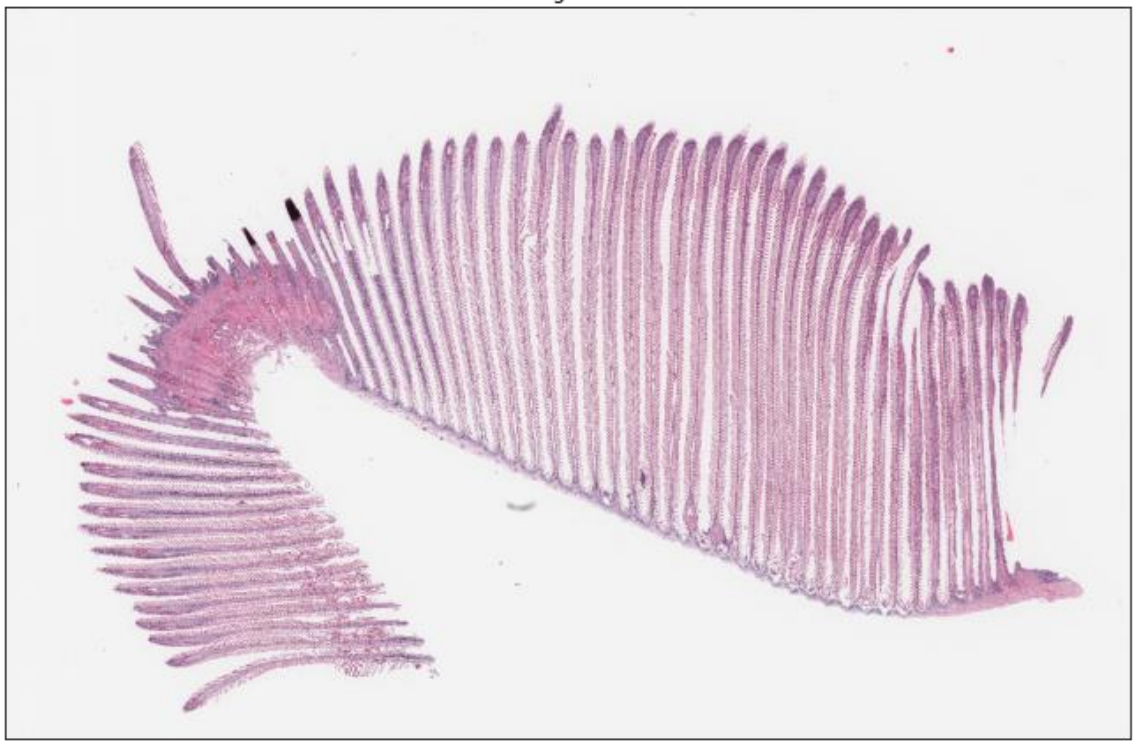
If the model is then given an image that it has not been trained to recognise, it produces a worse reconstruction.



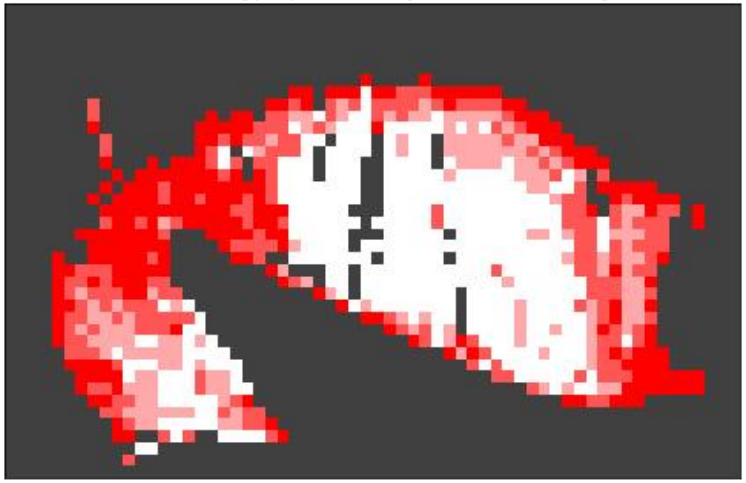
By comparing an image to the reconstruction, we get a value known as **Reconstruction Loss.**



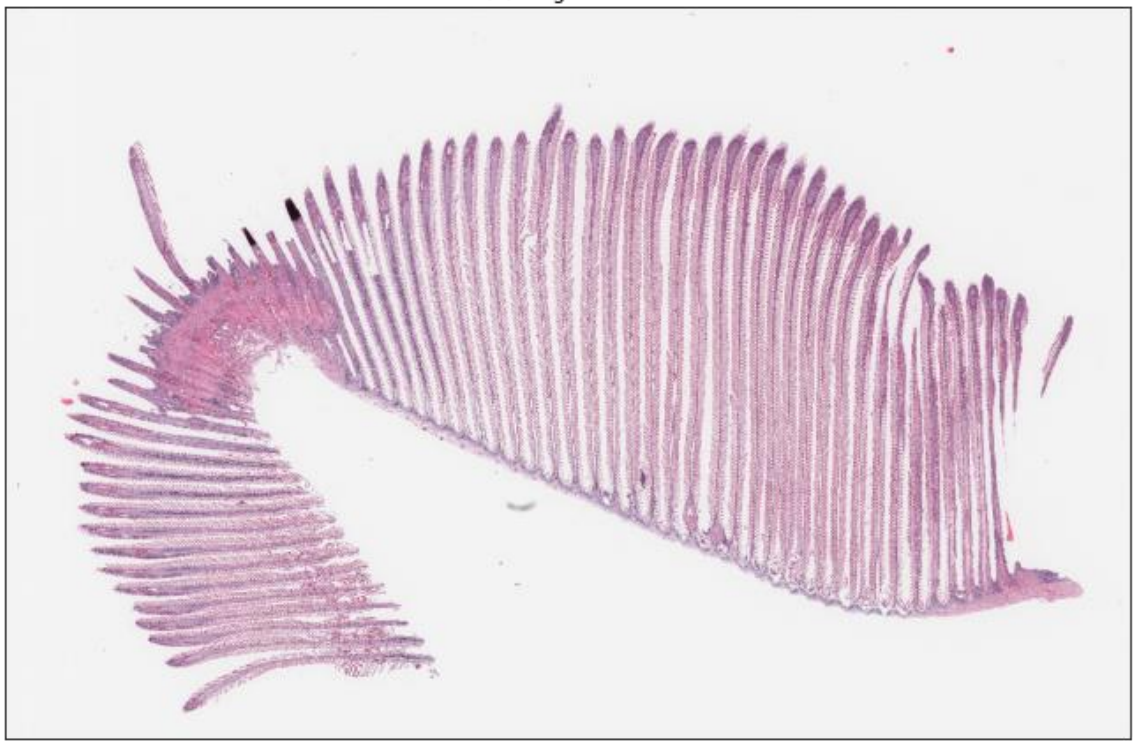
(a) Original WSI



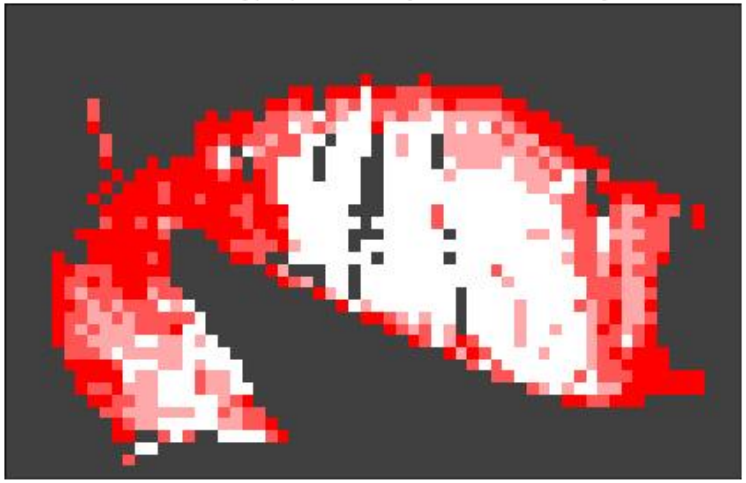
(b) Visualisation of Hyperplasia Analysis w/out Anomaly Detection



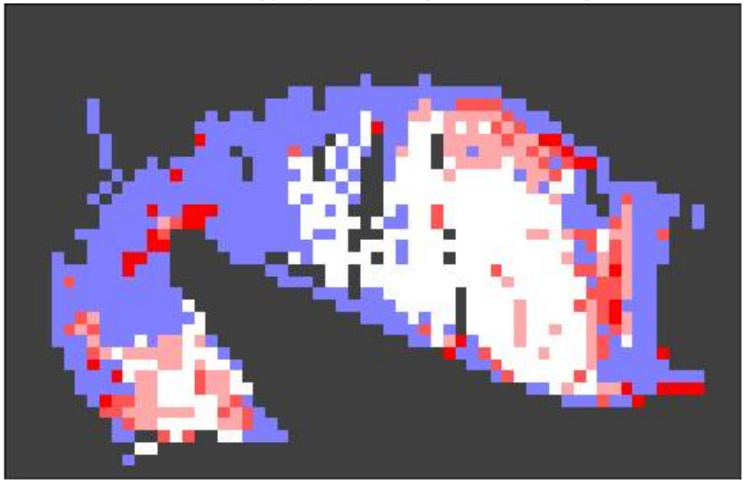
(a) Original WSI

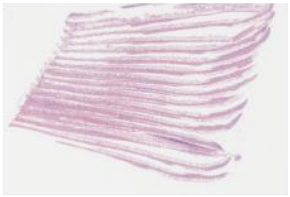


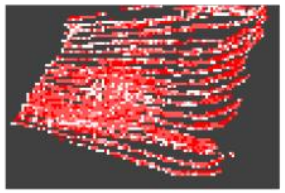
(b) Visualisation of Hyperplasia Analysis w/out Anomaly Detection

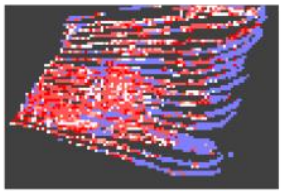


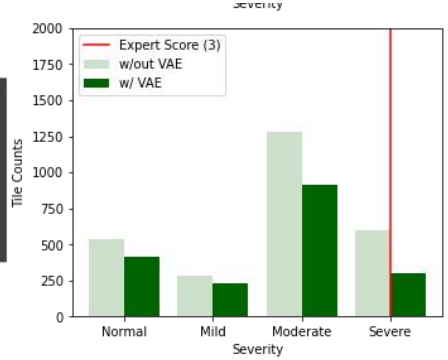
(c) Visualisation of Hyperplasia Analysis w/ Anomaly Detection

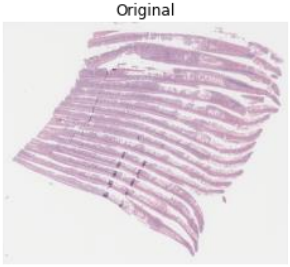


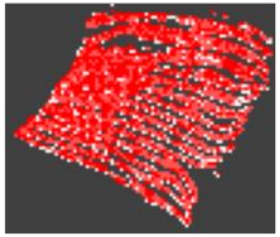

Original

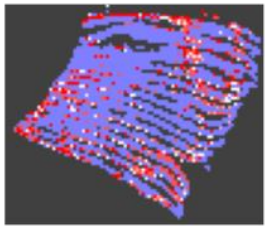

Visualisation w/out VAE

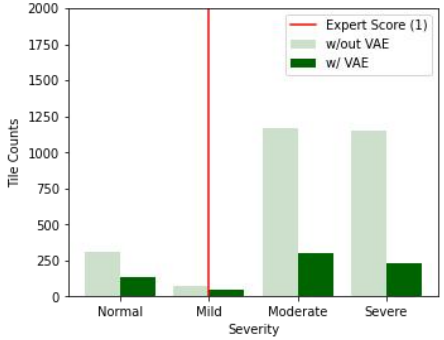

Visualisation w/ VAE


Expert Score (3)
w/out VAE
w/ VAE
Tile Counts
Normal
Mild
Moderate
Severe
Severity

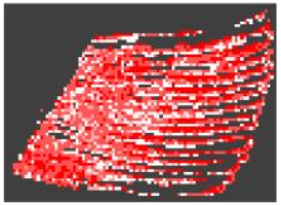

Original

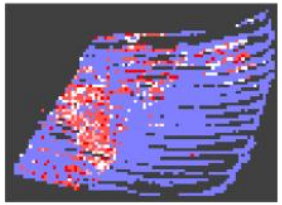

Visualisation w/out VAE

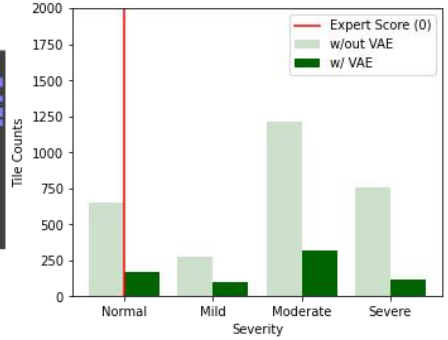

Visualisation w/ VAE


Expert Score (1)
w/out VAE
w/ VAE
Tile Counts
Normal
Mild
Moderate
Severe
Severity

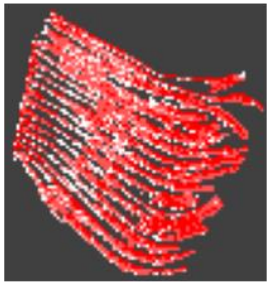

Original

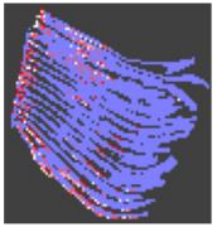

Visualisation w/out VAE

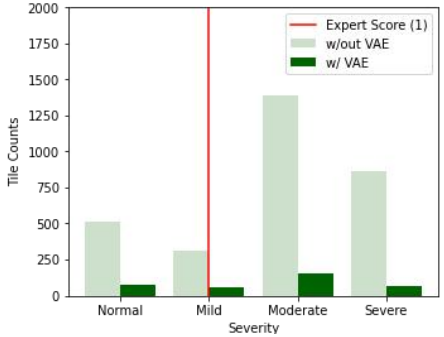

Visualisation w/ VAE


Expert Score (0)
w/out VAE
w/ VAE
Tile Counts
Normal
Mild
Moderate
Severe
Severity


Original


Visualisation w/out VAE


Visualisation w/ VAE


Expert Score (1)
w/out VAE
w/ VAE
Tile Counts
Normal
Mild
Moderate
Severe
Severity

How can we summarise this fine-grained analysis as a single score?

Does our analysis agree with expert opinion?

		Without Anomaly Detection						With Anomaly Detection					
		EWT-LP			ResNet18			EWT-LP			ResNet18		
ImageID	Expert Label	Mean	Mode	Median	Mean	Mode	Median	Mean	Mode	Median	Mean	Mode	Median
1	2	2	2	2	2	2	2	2	2	2	2	2	2
2	3	2	2	2	2	2	2	2	2	2	2	2	2
3	1	3	3	3	2	2	2	2	3	3	2	2	2
4	0	2	2	2	2	2	2	1	2	1	2	2	2
5	1	2	3	2	2	2	2	2	3	2	2	2	2
6	1	2	2	2	2	2	2	2	2	2	2	2	2
7	2	2	3	3	2	3	2	2	3	2	2	3	2
8	3	2	3	2	2	2	2	2	2	2	2	2	2
9	2	2	3	2	2	2	2	2	3	2	2	2	2
10	3	2	3	2	2	2	2	2	3	3	2	2	2
11	0	1	1	1	2	2	2	2	1	2	2	2	2
12	0	2	2	2	2	2	2	2	3	3	2	3	2
13	1	2	2	2	2	2	2	2	2	2	2	2	2
14	3	2	3	2	2	2	2	1	2	2	1	2	2
15	1	2	3	2	2	2	2	2	2	2	2	2	2
16	2	2	2	2	2	2	2	1	2	2	2	2	2
17	3	2	3	3	2	3	2	2	2	2	2	2	2
18	1	2	2	2	2	2	2	2	3	2	2	3	2
19	0	2	2	2	2	2	2	2	3	3	2	2	2
20	0	2	2	2	2	2	2	2	3	3	2	2	2

Table showing predicted labels based on mean, mode, and median averaging. Colour represents similarity to the expert label. Green: agreement between predicted and target. Yellow: Off-by-one between predicted and target. Orange: Off-by-two between predicted and target. Red: Off-by-three.

Final thoughts.

- There is significant scope for the application of advanced image processing techniques for **pathology** and l**esion-based analysis** in aquatic animal data.
- We have created a prototype **context-aware** tool for analysing hyperplasia in Atlantic Salmon gills WSIs that can assist histopathologists.
- Our context-aware approach and use of signal processing techniques means that the system is not a **“black box”** and can be interrogated/tested thoroughly.
- Our comparison to expert scores demonstrates a gap for more **fine-grained metrics** representing the **severity and distribution** of lesions.

SVIP SRUC UNIVERSITY of STIRLING

Any Questions?

a.f.carmichael@stir.ac.uk

BE THE DIFFERENCE